

WHAT IS CLAIMED IS:

1. A method of monitoring the operation of a deployed web site system, the method comprising:

5 (a) monitoring response times of a web site system as seen from multiple geographic locations, including locations that are geographically remote from each other and from the web site system;

10 (b) concurrently with (a), monitoring a plurality of server resource utilization parameters associated with the web site system from a computer that is local to the web site system; and

15 (c) displaying the response times and server resource utilization parameters as monitored in (a) and (b) over a selected time period as a function of time to allow an operator to evaluate whether a correlation exists between changes in the response times and changes in values of the plurality of server resource utilization parameters.

20 2. The method of Claim 1, wherein (a) comprises monitoring the response times from agent computers in at least some of the multiple geographic locations.

25 3. The method of Claim 1, wherein (a) comprises passively monitoring traffic resulting from actual web site users in at least some of the multiple geographic locations.

4. The method of Claim 1, wherein (a) comprises generating page requests from a data center, and sending the page requests to the web site system via Internet points of presence located in at least some of the multiple geographic locations.

5. The method of Claim 1, wherein (b) comprises monitoring at least one server resource utilization parameter of a web server.

25 6. The method of Claim 1, wherein (b) comprises monitoring at least one server resource utilization parameter of an application server.

7. The method of Claim 1, wherein (b) comprises monitoring at least one server resource utilization parameter of a database server.

30 8. The method of Claim 1, wherein (b) comprises monitoring at least one server resource utilization parameter of a network device.

9. The method of Claim 8, wherein the network device is a router.

10. The method of Claim 8, wherein the network device is a bridge.

11. The method of Claim 1, further comprising applying a statistical algorithm to a sequence of response time measurements resulting from (a) to automatically detect a degradation in performance.

5 12. The method of Claim 11, further comprising processing server resource utilization measurements resulting from (b) to identify at least one server resource parameter having a correlation with the degradation in performance.

13. A system for monitoring performance of a deployed transactional server, the system comprising:

10 a first agent configured to monitor a transactional server over a network, the first agent collecting performance data including response times of the transactional server;

15 a second agent configured to monitor server resource utilization of the transactional server, the second agent collecting data on one or more server resource utilization parameters, wherein the second agent monitors server resource utilization substantially concurrent with monitoring of the transactional server by the first agent; and

20 a report generating component configured to generate reports based at least on the data collected by the first and second agents, wherein at least one of the generated reports associates response times of the transactional server as monitored by the first agent with server resource utilization parameters as monitored by the second agent.

14. The system of Claim 13, wherein the first agent is configured to monitor network hop delays.

25 15. The system of Claim 13, wherein the first agent sends request messages to the transactional server to measure the response times.

16. The system of Claim 13, wherein the first agent passively monitors traffic between client computers and the transactional server to measure the response times.

30 17. The system of Claim 13, wherein the report generating component generates reports associating the response times with the server resource utilization

parameters by displaying the response times and the server resource utilization parameters on a time-synchronized graph to permit a human operator to determine whether a correlation exists between the response times and the server resource utilization parameters.

5        18.      The system of Claim 13, wherein the second agent is configured to monitor server resource utilization of a database server.

19.      The system of Claim 13, further comprising an analysis component that automatically detects correlations between response times and server resource utilization parameters.

10        20.      A method for monitoring the performance of a transactional server, the method comprising:

receiving performance data from a plurality of computers geographically distributed across a network, the plurality of computers executing transactions on a transactional server while monitoring associated response times;

15        receiving server resource utilization data from a computer that monitors server resource utilization of the transactional server during execution of the transactions by the plurality of computers; and

displaying the performance data in association with corresponding server resource data.

20        21.      The method of Claim 20, wherein the performance data includes time stamps for associating the performance data and the server resource utilization data.

22.      The method of Claim 20, wherein the server resource utilization data includes central process unit (CPU) utilization data associated with the transactional server.

25        23.      The method of Claim 20, wherein the server resource utilization data includes memory allocation data.

24.      The method of Claim 20, wherein the server resource utilization data includes at least one of the following: hits per second data, requests queued data, current connections data, connection attempts data, or disk utilization data.

25. A method of monitoring the operation of a deployed transactional server, the method comprising:

5 (a) monitoring response times of the transactional server as seen from multiple geographic locations, including locations that are geographically remote from each other and from the transactional server;

10 (b) concurrently with (a), monitoring a plurality of server resource utilization parameters associated with the transactional server; and

15 (c) displaying data indicative of whether a correlation exists between changes in the response times and changes in values of the plurality of server resource utilization parameters over time.

26. The method of Claim 25, wherein (c) comprises displaying response time data and server resource utilization data resulting from (a) and (b), respectively, on a like time scale to permit a human operator to determine whether the correlation exists.

27. The method of Claim 26, wherein (c) comprises displaying, for an operator-selected time window, a graph of the response times and a graph of at least one of the server resource utilization parameters.

28. The method of Claim 26, wherein (c) comprises analyzing response time data and server resource utilization data resulting from (a) and (b) with an automated analysis system that automatically detects correlations.

29. The method of Claim 25, wherein the transactional server is a web site system.